What is claimed is:

1. A method comprising:

modeling an audio-visual observation of a subject using a coupled Markov model to obtain an audio-visual model;

modeling a portion of the subject using an embedded Markov model to obtain a portion model; and

determining first and second likelihoods of identification based on the audio-visual model and the portion model.

2. The method of claim 1, wherein modeling the audio-visual observation comprises using a coupled hidden Markov model.

3. The method of claim 2, wherein the coupled hidden Markov model comprises a two-channel model, each channel having observation nodes coupled to backbone nodes via mixture nodes.

4. The method of claim 1, further comprising combining the first and second likelihoods of identification.

5. The method of claim 4, further comprising weighting the first and second likelihoods of identification.

6. The method of claim 1, wherein the portion of the subject comprises a mouth portion.

7. A method comprising:

recognizing a face of a subject from first entries in a database;

recognizing audio-visual speech of the subject from second entries in the database; and

identifying the subject based on recognizing the face and recognizing the audio-visual speech.

8. The method of claim 7, further comprising providing the subject access to a restricted area after identifying the subject.

9. The method of claim 7, wherein recognizing the face comprises modeling an image including the face using an embedded hidden Markov model.

10. The method of claim 9, further comprising obtaining observation vectors from a sampling window of the image.

11. The method of claim 10, wherein the observation vectors comprise discrete cosine transform coefficients.

12. The method of claim 7, wherein recognizing the face comprises performing a Viterbi decoding algorithm.

13. The method of claim 7, wherein recognizing the audio-visual speech further comprises detecting and tracking a mouth region using vector machine classifiers.

14. The method of claim 7, wherein recognizing the audio-visual speech comprises modeling an image and an audio sample using a coupled hidden Markov model.

15. The method of claim 7, further comprising combining results of recognizing the face and recognizing the audio-visual speech pattern according to a predetermined weighting to identify the subject.

16. A system comprising:

at least one capture device to capture audio-visual information from a subject;

a first storage device coupled to the at least one capture device to store code to enable the system to recognize a face of the subject from first entries in a

database, recognize audio-visual speech of the subject from second entries in the database, and identify the subject based on the face and the audio-visual speech; and

a processor coupled to the first storage to execute the code.

17. The system of claim 16, wherein the database is stored in the first storage device.

18. The system of claim 17, further comprising code that if executed enables the system to model an image including the face using an embedded hidden Markov model.

19. The system of claim 16, further comprising code that if executed enables the system to model an image and an audio sample using a coupled hidden Markov model.

20. An article comprising a machine-readable storage medium containing instructions that if executed enable a system to:

recognize a face of a subject from first entries in a database;

recognize audio-visual speech of the subject from second entries in the database; and

identify the subject based on recognizing the face and recognizing the audio-visual speech.

21. The article of claim 20, further comprising instructions that if executed enable the system to provide the subject access to a restricted area after the subject is identified.

22. The article of claim 20, further comprising instructions that if executed enable the system to model an image including the face using an embedded hidden Markov model.

23. The article of claim 20, further comprising instructions that if executed enable the system to model an image and an audio sample using a coupled hidden Markov model.